

A Gamma Distribution Hypothesis for Prime k -tuples

J. LaChapelle

Abstract

We conjecture average counting functions for prime k -tuples based on a gamma distribution hypothesis for prime powers. The conjecture is closely related to the Hardy-Littlewood conjecture for k -tuples. Possessing average counting functions along with the exact functions introduced in [6] allows to implicitly define pertinent k -tuple zeta functions. The k -tuple zeta functions in turn allow construction of k -tuple analogs of explicit formulae.

1 Introduction

The motivation and desire to better understand prime k -tuples hardly needs introduction. For a small sample of the literature see [1]–[5] and references therein. Of course there would be no suspense if there existed Euler products for k -tuples like that for single primes. Their absence seems to indicate that there do not exist generating Dirichlet series whose summands are completely multiplicative in these cases. It is fair to say that this is at the heart of the difficulty in generalizing the single prime case.

This paper hopefully takes a step toward better understanding the distribution of prime k -tuples. It is based on two ideas.

The first idea [6] uses the arithmetic function $\mu(n)\Lambda(n)/\log(n)$ to represent the exact prime counting function up to some cut-off $x \in \mathbb{R}_+$,

$$\pi(x) = - \sum_{n \leq x} \mu(n) \frac{\Lambda(n)}{\log(n)} .$$

This simple representation can be readily extended to prime k -tuples determined by an admissible set $\mathcal{H}_k = \{0, h_2, \dots, h_k\}$:

$$\pi_{(k)}(x) = (-1)^k \sum_{n \leq x} \mu(n) \cdots \mu(n + h_k) \frac{\Lambda(n)}{\log(n)} \cdots \frac{\Lambda(n + h_k)}{\log(n + h_k)}$$

The second idea is the gamma distribution hypothesis: *prime powers* are random variables on the *positive reals*, and counting them is a random process following a non-homogenous gamma distribution [7]. The resulting probability model — a non-homogenous Poisson process — yields quite accurate average counting functions associated with the primes.

Therefore it is natural to generalize to prime k -tuples to test the hypothesis. The obvious tack is to consider a joint gamma distribution on \mathbb{R}_+^k . But, in light of the exact k -tuple counting function, the counting is modeled by a probability distribution along a ray $\mathbf{r}_k \in \mathbb{R}_+^k$ determined by an admissible set $\mathcal{H}_k = \{0, h_2, \dots, h_k\}$. Taking this into account leads to an ansatz for the density of k -tuples of prime powers along \mathbf{r}_k up to some cut-off \mathbf{x}

$$P_{(k)}(n; \mathbf{x}) := \frac{(-1)^{n-1}}{n!} \int_0^{\mathbf{x}} (\log(r))^{n-1} \cdots (\log(r + h_k))^{n-1} dr$$

where the integral is defined by the Cauchy principal value.

There is no reason to expect the probability measure on \mathbf{r}_k to coincide with the probability measure on \mathbb{R}_+ for the single prime case, and we will argue that

$$dr = C_{(k)}(\mathbf{x}) dt$$

where $C_{(k)}(\mathbf{x})$ is a suitable normalization defined in (A.8). For admissible k -tuples, the ansatz leads to accurate counting functions because asymptotically $C_{(k)}(\mathbf{x}) \sim C_{(k)}$ where $C_{(k)}$ is the appropriate singular series, i.e. the prime k -tuple constant.

But the enumeration is secondary. The primary goal is to extract information about prime k -tuple distributions, which means we need to discover pertinent k -tuple zeta functions¹ implicitly defined by

$$\log(\zeta_{(k)}(s)) := \sum_{n=1}^{\infty} \frac{\Lambda_{(k)}(n)}{\log_{(k)}(n) n_{(k)}^s}$$

where $\Lambda_{(k)}(n) := \Lambda(n) \cdots \Lambda(n + h_k)$, similarly $\log_{(k)}(n) := \log(n) \cdots \log(n + h_k)$, and $n_{(k)}$ denotes the geometric mean of $n_k := (n, n + h_2, \dots, n + h_k)$. These log-zeta functions are just what one would guess from the structure of the k -tuple analogs of the first Chebyshev function [6].

Unfortunately we haven't found an explicit representation of $\zeta_{(k)}(s)$ that would allow the prime k -tuple issue to be settled. But in the final section we motivate the fairly obvious conjecture that

$$\zeta_{(k)}(s) = \sum_{n_{(k)}} \frac{1}{n_{(k)}^s} = \sum_{n=1}^{\infty} \frac{1}{n_{(k)}^s}.$$

The first sum is over the geometric means of points along a ray in the pair-wise coprime k -lattice denoted $\mathfrak{N}_+^k \subset \mathbb{N}_+^k$.

This point bears repeating: Possessing exact counting functions (in terms of standard arithmetic functions) and a model probability distribution facilitates constructing k -tuple zeta functions and, subsequently, explicit integral representations of certain counting functions.² This opens the possibility to attack the problem of prime k -tuple distributions using more-or-less elementary methods borrowed from the single prime case.

¹The zeta functions implicitly defined by the log-zeta functions will not be studied here, but they are of obvious interest.

²We emphasize that our explicit formulae are left as integral representations. Since we do not determine the complex analytic properties of the k -tuple zeta functions, we cannot express the integrals in terms of residues.

2 Prime k -tuple conjecture

According to [7], events along a directed graph can be modeled by a suitable gamma distribution. For prime k -tuples the events occur along a ray $\mathbf{r}_k \in \mathbb{R}_+^k$ determined by an admissible set $\mathcal{H}_k = \{0, h_2, \dots, h_k\}$. We are counting prime-power events up to some cut-off \mathbf{x} , and according to the gamma hypothesis this is a scaled Poisson process.

To learn how to apply the gamma hypothesis for k -tuples, let's briefly review the single prime case. Consider a Poisson process with trivial mean. In this case, counting corresponds to the observation of integers because the events are evenly distributed with unit density. Accordingly, the trivial gamma distribution on \mathbb{R}_+ and its associated Poisson process yield a model of the positive integers \mathbb{Z}_+ as the cut-off $\mathbf{x} \rightarrow \infty$, because they are in one-to-one correspondence with the positive natural numbers \mathbb{N}_+ .

Less trivially, the gamma hypothesis posits the expected number of prime-power events along \mathbb{R}_+ is given by

$$\overline{N(\mathbf{x})} = \sum_{n=1}^{\infty} \frac{(-1)^1}{n!} \gamma(n, -\log(\mathbf{x})) . \quad (2.1)$$

This has the expected form of a scaled Poisson expectation [7]; suggesting we write the incomplete gamma function as an integral in order to infer the associated prime-power probability distribution on \mathbb{R}_+ :

$$\begin{aligned} \overline{N(\mathbf{x})} &= \sum_{n=1}^{\infty} \frac{(-1)^1}{n!} \int d\gamma(n, -\log(\mathbf{x})) \\ &= \sum_{n=1}^{\infty} \frac{1}{n!} \int (-\log(\mathbf{x}))^{n-1} dx \\ &= - \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \int_0^{\mathbf{x}} (\log(t))^{n-1} dt . \end{aligned} \quad (2.2)$$

where the integral in the third line is taken as the Cauchy principal value. We infer the probability distribution of prime powers goes like $\log(t)^{-1}$. Interchanging the sum and integral yields³ $\overline{N(\mathbf{x})} = \text{Ei}(\log(\mathbf{x})) - \log(\log(\mathbf{x}))$.

Return now to the general case. The integral in (2.2) becomes a multiple integral on \mathbb{R}_+^k , and the probability distribution will be a k -fold product of logarithms restricted to the appropriate ray determined by \mathcal{H}_k . Consequently, for the general k -tuple case the mean number of prime-power k -tuple events is expected to be

$$\begin{aligned} \overline{N_{(k)}(\mathbf{x})} &= - \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \int_0^{\mathbf{x}} (\log(r))^{n-1} \cdots (\log(r + h_k))^{n-1} dr \\ &\propto - \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \int_0^{\mathbf{x}} (\log(t))^{n-1} \cdots (\log(t + h_k))^{n-1} dt \end{aligned} \quad (2.3)$$

³We use $\text{Ei}(\log(\mathbf{x}))$ instead of $\text{li}(\mathbf{x})$ to remind that the gamma hypothesis applies to the more general case of complex cut-off $\mathbf{x} \in \mathbb{C}_+$.

where $r \in \mathbf{r}_k$ and $t \in \mathbb{R}_+$. In Appendix A we argue the ratio of the two integrators is given by

$$C_{(k)}(\mathbf{x}) := \prod_{p \leq \mathbf{x}_{(k)}^k} \left(1 - \frac{\nu_p(\mathcal{H}_k)}{p}\right) \prod_p \left(1 - \frac{1}{p}\right)^{-k}. \quad (2.4)$$

This leads to

Ansatz 2.1 *Let $\mathcal{H}_k = \{0, h_2, \dots, h_k\}$ be admissible. The expected number of events associated with counting admissible prime-power k -tuples up to some cutoff \mathbf{x} is*

$$\overline{N_{(k)}(\mathbf{x})} = \sum_{n=1}^{\infty} P_{(k)}(n; \mathbf{x}) := - \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} I_{(k)}(n; \mathbf{x}) \quad (2.5)$$

where

$$\begin{aligned} I_{(k)}(n; \mathbf{x}) &:= \int_0^{\mathbf{x}} (\log(r))^{n-1} \cdots (\log(r + h_k))^{n-1} dr \\ &= C_{(k)}(\mathbf{x}) \int_0^{\mathbf{x}} (\log(t))^{n-1} \cdots (\log(t + h_k))^{n-1} dt \\ &=: C_{(k)}(\mathbf{x}) \int_0^{\mathbf{x}} \log_{(k)}^{n-1}(t) dt \end{aligned} \quad (2.6)$$

with the integral defined by the principal value and $C_{(k)}(\mathbf{x}) \sim C_{(k)}$ the singular series.

The analysis in [7] suggests that

$$\overline{N_{(k)}(\mathbf{x})} \approx \sum_{n \leq \mathbf{x}} \lambda_{(k)}(n) - \sum_{n | \mathbf{x}} \lambda_{(k)}(n). \quad (2.7)$$

where

$$\begin{aligned} \lambda_{(k)}(n) &:= \frac{\Lambda(n) \cdots \Lambda(n + h_k)}{\log(n) \cdots \log(n + h_k)} \\ &=: \frac{\Lambda_{(k)}(n)}{\log_{(k)}(n)}. \end{aligned} \quad (2.8)$$

Moreover, since the sum and integral can be interchanged,

$$\overline{N_{(k)}(\mathbf{x})} = C_{(k)}(\mathbf{x}) \int_0^{\mathbf{x}} \frac{1}{\log_{(k)}(t)} dt - \text{small remainder} \quad (2.9)$$

for all $k \in \mathbb{N}_+$. Again, [7] suggests the small remainder term represents $\sum_{n | \mathbf{x}} \lambda_{(k)}(n)$ while

$$\begin{aligned} \overline{J_{(k)}(\mathbf{x})} &:= C_{(k)}(\mathbf{x}) \int_0^{\mathbf{x}} \frac{1}{\log_{(k)}(t)} dt =: C_{(k)}(\mathbf{x}) \text{Ei}_{(k)}(\log(\mathbf{x})) \\ &\approx \sum_{n \leq \mathbf{x}} \lambda_{(k)}(n) \end{aligned} \quad (2.10)$$

is the average k -tuple analog of Riemann's counting function.

In particular, for prime doubles

$$\begin{aligned}\overline{N_{(2)}(x)} &= C_{(2)}(x) \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n!} \int_0^x (\log(t) \log(t + h_{2i}))^{n-1} dt \\ &=: \overline{J_{(2)}(x)} - \overline{\omega_{(2)}(x)}.\end{aligned}\tag{2.11}$$

where $C_{(2)} = C_2 \prod_{p|i} \frac{p-1}{p-2}$ with C_2 the twin prime constant.

Given this heuristic motivation, we conjecture:

Conjecture 2.1 *Given an admissible \mathcal{H}_k , the average number of admissible prime k -tuples up to some cut-off integer x is*

$$\overline{\pi_{(k)}(x)} = \sum_{m=1}^{\infty} \frac{\mu(m)}{m} \overline{J_{(k)}(x^{1/m})}.\tag{2.12}$$

The reader is encouraged to numerically explore (2.12). Note that, whereas the Hardy-Littlewood conjecture is asymptotic, (2.12) holds for all $x > 2$. The difference between predicted counts for Conjecture 2.1 and the Hardy-Littlewood conjecture is especially stark when $x \ll h_k$.

Analogous reasoning helps to define the average prime double Chebyshev function:

Definition 2.1

$$\begin{aligned}\overline{\psi_{(2)}(x)} &:= C_{(2)}(x) \int_0^x \frac{\log(t_{(2)})}{\log(t) \log(t + h_{2i})} dt \\ &\approx \sum_{n \leq x} \frac{\Lambda_{(2)}(n)}{\log_{(2)}(n)} \log(n_{(2)})\end{aligned}\tag{2.13}$$

where $t_{(2)} := (t(t + h_2))^{1/2}$ is the geometric mean of $(t, t + h_2)$.

This has obvious extensions to higher k -tuples:

Definition 2.2

$$\begin{aligned}\overline{\psi_{(k)}(x)} &:= C_{(k)}(x) \int_0^x \frac{\log(t_{(k)})}{\log_{(k)}(t)} dt \\ &\approx \sum_{n \leq x} \frac{\Lambda_{(k)}(n)}{\log_{(k)}(n)} \log(n_{(k)})\end{aligned}\tag{2.14}$$

Conjecture 2.2

$$\overline{\theta_{(2)}(x)} = \sum_{m=1}^{\infty} \mu(m) \overline{\psi_{(2)}(x^{1/m})}.\tag{2.15}$$

Note that $\overline{\psi_{(2)}(x)} = C_{(2)}(x)(\text{Ei}(\log(x)) + \text{Ei}(\log(x + h_2)) - \text{Ei}(\log(h_2)))/2$ follows from (2.13). Hence $\overline{\theta_{(2)}(x)} \sim C_{(2)}(x/\log(x))$ which is consistent with the Hardy-Littlewood twin prime conjecture.

3 Explicit formulae

Having both exact and average summatory functions allows to deduce associated k -tuple zeta functions and subsequent explicit formulae. Here we will confine attention to prime doubles but indicate the generalization to higher k -tuples.

Define the prime double zeta function *implicitly* by

Definition 3.1

$$\log(\zeta_{(2)}(s)) := \sum_{n=1}^{\infty} \frac{\lambda_{(2)}(n)}{n^{s/2}(n+2i)^{s/2}} = \sum_{n=1}^{\infty} \frac{\Lambda_{(2)}(n)}{\log_{(2)}(n) n_{(2)}^s}, \quad \Re(s) > 1. \quad (3.1)$$

It follows that

$$\log'(\zeta_{(2)}(s)) = \frac{\zeta'_{(2)}(s)}{\zeta_{(2)}(s)} = - \sum_{n=1}^{\infty} \frac{\Lambda_{(2)}(n)}{\log_{(2)}(n) n_{(2)}^s} \log(n_{(2)}). \quad (3.2)$$

Using this log-zeta function, along with the gamma hierarchy from [7] as a guide, we construct an explicit formula for

$$\psi_{(2)}(\mathbf{x}) = \sum_{n \leq \mathbf{x}} \frac{\Lambda_{(2)}(n)}{\log_{(2)}(n)} \log(n_{(2)}). \quad (3.3)$$

Proposition 3.1 *Put $\tilde{\mathbf{x}} = \mathbf{x} + \epsilon$ with $\mathbf{x} \in \mathbb{N}_+$ and $0 < \epsilon < 1$. Let σ_a be the abscissa of absolute convergence of $\sum_{n=1}^{\infty} \frac{\lambda_{(2)}(n) \log(n_{(2)})}{n^{s/2}(n+2i)^{s/2}}$. Then, for $c > \sigma_a$,*

$$\begin{aligned} \psi_{(2)}(\mathbf{x}) &= - \lim_{\epsilon \rightarrow 0} \lim_{T \rightarrow \infty} \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \Gamma(0, -\log(\tilde{\mathbf{x}}_{(2)}^s)) d \log'(\zeta_{(2)}(s)), \quad c > \sigma_a \\ &= \lim_{\epsilon \rightarrow 0} \lim_{T \rightarrow \infty} \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \text{Ei}(\log(\tilde{\mathbf{x}}_{(2)}^s)) d \log'(\zeta_{(2)}(s)), \quad c > \sigma_a \\ &= \sum_{n \leq \mathbf{x}} \frac{\Lambda_{(2)}(n)}{\log_{(2)}(n)} \log(n_{(2)}). \end{aligned} \quad (3.4)$$

Proof: First integrate (3.4) by parts. The boundary term does not contribute because i) a comparison test yields a finite σ_a (in fact $\sigma_a = 1$) so $\lim_{t \rightarrow \infty} |\log'(\zeta_{(2)}(c+it))| < \infty$ for $c > \sigma_a$; and ii) $\lim_{t \rightarrow \infty} |\text{Ei}(\log(\tilde{\mathbf{x}}_{(2)}^s))| = 0$ since

$$\begin{aligned} \lim_{t \rightarrow \infty} |\text{Ei}(\log(\mathbf{x}^{(c+it)}))| &= \lim_{t \rightarrow \infty} \left| \frac{\mathbf{x}^{(c+it)}}{(c+it) \log(\mathbf{x})} \left(1 + O\left(\frac{1}{(c+it) \log(\mathbf{x})}\right) \right) \right| \\ &\leq \frac{\mathbf{x}^c}{\log(\mathbf{x})} \lim_{t \rightarrow \infty} \left| \frac{1}{(c+it)} \left(1 + O\left(\frac{1}{(c+it)}\right) \right) \right| = 0. \end{aligned} \quad (3.5)$$

Next, following standard arguments, use the truncating integral

Lemma 3.1

$$\frac{1}{2\pi i} \int_{c-iT}^{c+iT} \frac{x^s}{s} ds = \begin{cases} 1 + O\left(\frac{x^c}{T \log(x)}\right) & x > 1 \\ O\left(\frac{x^c}{T \log(x)}\right) & 0 < x < 1 \end{cases}. \quad (3.6)$$

proof: We include the proof for completeness. For $x > 1$ integrate over a rectangle with left edge $(L - iT, L + iT)$ such that $L < c$. We have

$$\lim_{L \rightarrow -\infty} \left| \int_{L-iT}^{L+iT} \frac{x^s}{s} ds \right| \leq \lim_{L \rightarrow -\infty} \int_{-T}^T \frac{x^L}{|L + it|} dt < \lim_{L \rightarrow -\infty} \frac{T x^L}{L} = 0. \quad (3.7)$$

The top and bottom contribute

$$\begin{aligned} \left| \int_{-\infty \pm iT}^{c \pm iT} \frac{x^s}{s} ds \right| &\leq \int_{-\infty}^0 \frac{-x^{c-r}}{|(c-r) \pm iT|} dr \\ &= x^c \int_{-\infty}^0 \frac{-x^{-r}}{|(c-r) \pm iT|} dr \\ &< x^c \int_{-\infty}^0 \frac{-x^{-r}}{T} dr \\ &= \frac{x^c}{T \log(x)}. \end{aligned} \quad (3.8)$$

Finally, the pole at $s = 0$ contributes $\text{Res} = 1$.

Now, for $x < 1$ integrate over the right edge $(R - iT, R + iT)$ with $c < R$. Then

$$\lim_{R \rightarrow \infty} \left| \int_{R-iT}^{R+iT} \frac{x^s}{s} ds \right| \leq \lim_{R \rightarrow \infty} \int_{-T}^T \frac{e^{-R|\log(x)|}}{|R + it|} dt < \lim_{R \rightarrow \infty} \frac{T e^{-R|\log(x)|}}{R} = 0. \quad (3.9)$$

The top and bottom contribute the same order as for $x > 1$, so the well-known lemma is established. \square

Hence, for $c > \sigma_a$,

$$\begin{aligned} -\lim_{\epsilon \rightarrow 0} \lim_{T \rightarrow \infty} \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \log'(\zeta_{(2)}(s)) \frac{\widetilde{x_{(2)}}^s}{s} ds \\ = \lim_{\epsilon \rightarrow 0} \lim_{T \rightarrow \infty} \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \sum_{n=1}^{\infty} \frac{\lambda_{(2)}(n) \log(n_{(2)})}{n^{s/2} (n+2i)^{s/2}} \frac{\widetilde{x_{(2)}}^s}{s} ds \\ = \lim_{\epsilon \rightarrow 0} \lim_{T \rightarrow \infty} \sum_{n=1}^{\infty} \frac{\Lambda_{(2)}(n)}{\log_{(2)}(n)} \log(n_{(2)}) \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \frac{\left(\widetilde{x}^{1/2} (\widetilde{x}+2i)^{1/2}\right)^s}{n^{s/2} (n+2i)^{s/2}} \frac{ds}{s} \\ = \lim_{\epsilon \rightarrow 0} \sum_{n \leq [\widetilde{x}]} \frac{\Lambda_{(2)}(n)}{\log_{(2)}(n)} \log(n_{(2)}) \\ = \sum_{n \leq x} \frac{\Lambda_{(2)}(n)}{\log_{(2)}(n)} \log(n_{(2)}) \end{aligned} \quad (3.10)$$

where the third equality follows from the lemma. (Justifying the interchange of the sum and integral is straightforward, and interchange of the T -limit and sum is allowed because the summand contains $O(n^{-c})$ with $c > 1$.) \square

Clearly this result only has teeth if one possesses an explicit representation of $\zeta_{(2)}(s)$. *But if* a suitable representation of $\zeta_{(2)}(s)$ can be found and it enjoys analytic properties similar to $\zeta(s)$, then we might expect

$$\begin{aligned}\psi_{(2)}(\mathbf{x}) &= \lim_{T \rightarrow \infty} \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \text{Ei}(\log(\mathbf{x}_{(2)}^s)) d \log'(\zeta_{(2)}(s)) \\ &= \lim_{T \rightarrow \infty} \frac{1}{2\pi i} \int_{c-iT}^{c+iT} \text{Ei}(\log(\mathbf{x}_{(2)}^s)) \log''(\zeta_{(2)}(s)) ds\end{aligned}\quad (3.11)$$

would lead to something like

$$\psi_{(2)}(\mathbf{x}) \sim C_{(2)} \text{Ei}(\log(\mathbf{x})) - C_{(2)} \sum_{\rho_{(2)}} \text{Ei}(\log(\mathbf{x}^{\rho_{(2)}})) + \text{small terms} \quad (3.12)$$

where the sum would include conjugate pairs of nontrivial zeros of $\zeta_{(2)}(s)$. *And if* nontrivial zeros of $\zeta_{(2)}(s)$ are confined within its critical strip, then the same proof strategy used for the PNT would appear to apply to prime doubles. Evidently, if this scenario plays out the prime double constant $C_{(2)}$ will have to come from $\log''(\zeta_{(2)}(s))$.

Remark that for higher k -tuples one should define

Definition 3.2

$$\overline{\varphi_{(k)}(\mathbf{x})} := C_{(k)}(\mathbf{x}) \int_0^{\mathbf{x}} \frac{\log^{k-1}(t_{(k)})}{\log_{(k)}(t)} dt. \quad (3.13)$$

and the k -tuple log-zeta function

Definition 3.3

$$\log(\zeta_{(k)}(s)) := \sum_{n=1}^{\infty} \frac{\lambda_{(k)}(n)}{n_{(k)}^s} = \sum_{n=1}^{\infty} \frac{\Lambda_{(k)}(n)}{\log_{(k)}(n) n_{(k)}^s} \quad (3.14)$$

so that

$$\log^{(k-1)'}(\zeta_{(k)}(s)) = (-1)^{k-1} \sum_{n=1}^{\infty} \frac{\Lambda_{(k)}(n)}{\log_{(k)}(n) n_{(k)}^s} \log^{k-1}(n_{(k)}) . \quad (3.15)$$

Then to construct an explicit formula at level k , consider

$$\lim_{T \rightarrow \infty} \frac{(-1)^{k-1}}{2\pi i} \int_{c-iT}^{c+iT} \Gamma(0, -\log(\mathbf{x}_{(k)}^s)) d \log^{(k-1)'}(\zeta_{(k)}(s)) . \quad (3.16)$$

Assuming integration by parts to be valid, Perron's formula would apply yielding

$$\varphi_{(k)}(\mathbf{x}) = \sum_{n \leq \mathbf{x}} \frac{\Lambda_{(k)}(n)}{\log_{(k)}(n)} \log^{k-1}(n_{(k)}) . \quad (3.17)$$

Finally, assuming favorable analytic properties for $\zeta_{(k)}(s)$ similar to $\zeta(s)$, one would expect to find $\varphi_{(k)}(\mathbf{x}) \sim C_{(k)} \text{Ei}(\log(\mathbf{x}))$.

4 Searching for $\zeta_{(k)}(s)$

The manipulations in the previous section point to a possible representation for $\zeta_{(k)}(s)$ which we formulate as another conjecture. First note that

$$\begin{aligned}
\log(\zeta_{(2)}(s)) &> \sum_{\mathfrak{p}_2 \in \mathfrak{P}_2} \sum_{\omega, \omega': \omega = \omega'} \frac{1}{\omega p^{\omega s/2}} \frac{1}{\omega' (p^\omega + 2i)^{\omega' s/2}}, \quad \Re(s) > 1 \\
&> \sum_{\mathfrak{p}_2 \in \mathfrak{P}_2} \sum_{\omega, \omega': \omega = \omega'} \frac{1}{\omega (p^{\omega'})^{\omega s/2}} \frac{1}{\omega' ((p + 2i)^\omega)^{\omega' s/2}} \\
&= \sum_{\mathfrak{p}_2 \in \mathfrak{P}_2} \sum_{\omega^2} \frac{1}{\omega^2 p_{(2)}^{\omega^2}} \\
&= - \sum_{\mathfrak{p}_2 \in \mathfrak{P}_2} \log(1 - p_{(2)}^{-s}). \tag{4.1}
\end{aligned}$$

This suggests to define $Z_{(k)}(s) := \prod_{\mathfrak{p}_k \in \mathfrak{P}_k} (1 - p_{(k)}^{-s})^{-1}$ which motivates

Conjecture 4.1 *Let \mathfrak{r}_k be a ray in the pair-wise coprime k -lattice \mathfrak{N}_+^k and $n_{(k)}$ the geometric mean of a point $n_k = (n, n + h_2, \dots, n + h_k) \in \mathfrak{r}_k$. Then*

$$\zeta_{(k)}(s) = \sum_{n_{(k)}} \frac{1}{n_{(k)}^s} = \sum_{n=1}^{\infty} \frac{1}{n_{(k)}^s} \tag{4.2}$$

where the first sum is over the geometric means of all points along $\mathfrak{r}_k \in \mathfrak{N}_+^k$.

If the conjecture is correct, the k -tuple zeta function is evidently a restriction of the multiple zeta function which motivates

Conjecture 4.2 *$\zeta_{(k)}(s)$ is meromorphic on \mathbb{C} , and the singular part of $(-1)^k \log^{(k)'}(\zeta_{(k)}(s))$ is given by*

$$\frac{1}{s-1} \frac{(-1)^k}{2\pi i} \oint \log^{(k)'}(\zeta_{(k)}(s)) ds = \frac{C_{(k)}}{s-1}. \tag{4.3}$$

This conjecture is equivalent to the Hardy-Littlewood prime k -tuple conjecture in the following sense. If we believe the gamma hypothesis, then being a sum of $1/n_{(k)}^s$ along a ray in the pair-wise coprime k -lattice *strongly suggests* $\zeta_{(k)}(s)$ has a first order pole at $s = 1$ and there are no other poles, while its zeros are determined by conspiring projections of almost periodic exponentials.

Together with the explicit formula at level k the conjecture implies

$$\begin{aligned}
\overline{\varphi_{(k)}(\mathbf{x})} &= \lim_{T \rightarrow \infty} \frac{(-1)^k}{2\pi i} \int_{c-iT}^{c+iT} \text{Ei}(\log(\mathbf{x}_{(k)}^s)) \log^{(k)'}(\zeta_{(k)}(s)) ds \Big|_{s=1} \\
&\sim C_{(k)} \text{Ei}(\log(\mathbf{x})). \tag{4.4}
\end{aligned}$$

Equivalently,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} \frac{\Lambda_{(k)}(n)}{\log_{(k)}(n)} \log^k(n_{(k)}) = C_{(k)} . \quad (4.5)$$

Of course, possessing poles and zeros of $\zeta_{(k)}(s)$ would be tantamount to evaluating the exact summatory functions. And their singular part would presumably furnish the prime k -tuple constants. The goal would be to express the integral in Proposition 3.1 as a sum over k -tuple zeta residues in the usual way; which would presumably verify Conjecture 2.1 and validate the gamma hypothesis.

Acknowledgments: I thank A. Granville for offering some helpful advice.

A The k -tuple normalization

In this appendix we argue that $C_{(k)}(\mathbf{x})$ is asymptotically the prime k -tuple constant. Since the argument is based on the gamma distribution hypothesis, it is very close to the standard probability argument. However, the details are a bit different since the counting occurs in the pair-wise coprime lattice.

According to the probability model [7] for the case $k = 1$, there is a distinction between the positive integer lattice \mathbb{Z}_+ and the strictly positive natural numbers \mathbb{N}_+ that characterize the counting process associated with the trivial gamma distribution on \mathbb{R}_+ . However, the trivial gamma distribution yields a Poisson process with unit integer density, and this allows for a meaningful identification.

But in general the probability model is a counting process along a ray \mathbf{r}_k in the k -lattice, and there is no guarantee that the trivial gamma distribution along \mathbf{r}_k leads to a unit integer density. Accordingly, in virtue of restricting to the pair-wise coprime k -lattice, the integer density may depend on \mathcal{H}_k and \mathbf{x} ; and it is necessary to renormalize the probability distribution along \mathbf{r}_k if one wants to compare counting processes and maintain the natural identification of \mathbb{N}_+ and \mathbb{Z}_+ .

In other words, comparing exact counting functions to average counting functions in the single prime case, we can interpret

$$\sum_{n \leq \mathbf{x}} \frac{\Lambda(n)}{\log(n)} \longrightarrow \int_0^{\mathbf{x}} \frac{dt}{\log(t)} \quad (A.1)$$

as a representation of the averaging process, and then dt is the integrator associated with the probability measure of prime powers — a kind of smoothed $\Lambda(n)$. For the general case we have

$$\sum_{n \leq \mathbf{x}} \frac{\Lambda_{(k)}(n)}{\log_{(k)}(n)} \longrightarrow \int_0^{\mathbf{x}} \frac{dr}{\log_{(k)}(r)} . \quad (A.2)$$

But now dr encodes both the ‘smoothed’ $\Lambda_{(k)}(n)$ and the density of integers along \mathbf{r}_k .

In order to put the density of integers along \mathbb{R}_+ and \mathfrak{r}_k on equal footing, we use the gamma hypothesis for prime powers together with the fundamental theorem of arithmetic. Evidently, it suffices to deduce their ratio for the particular case of counting prime powers. So the task is to determine the prime-power density along \mathfrak{r}_k relative to the prime-power density along \mathbb{R}_+ .

We are looking for the probability that the k -th power of the geometric mean of a point $n_k = (n, n + h_2, \dots, n + h_k) \in \mathfrak{r}_k$ is coprime to some prime, since this is a necessary condition for $n_{(k)}^k$ to be a prime power. Our main tool is a theorem by Tóth [8]: Let $k, m, u \geq 1$ and

$$P_k^{(u)}(m) = \sum_{\substack{1 \leq a_1, \dots, a_k \leq m \\ (a_i, a_j) = 1, i \neq j \\ (a_i, u) = 1}} 1 \quad (\text{A.3})$$

be the number of k -tuples (a_1, \dots, a_k) on the pair-wise coprime lattice with $1 \leq a_i \leq m$ and $(a_i, u) = 1$ for all $i \in \{1, \dots, k\}$.

Theorem A.1 ([8]) *For a fixed $k \geq 1$, we have uniformly for $m, u \geq 1$,*

$$P_k^{(u)}(m) = A_k f_k(u) m^k + O(\theta(u) m^{k-1} \log^{k-1}(m)) \quad (\text{A.4})$$

where

$$\begin{aligned} A_k &= \prod_{p'} \left(1 - \frac{1}{p'}\right)^{k-1} \left(1 + \frac{k-1}{p'}\right) \\ f_k(u) &= \prod_{p' | u} \left(1 - \frac{k}{p' + k - 1}\right) \end{aligned}$$

and $\theta(u)$ is the number of squarefree divisors of u .

Restrict $P_k^{(u)}(m)$ to the ray \mathfrak{r}_k and choose $m > n + h_k$. Then the density of points along \mathfrak{r}_k that are coprime to a given prime (or prime power) p is

$$D_k^{(p)}(m) := P_k^{(p)}(m)/m^k = A_k f_k(p) + O(\log^{k-1}(m)/m) . \quad (\text{A.5})$$

Of course m is automatically coprime to all primes $p > m$. Consequently, the density of prime powers along \mathfrak{r}_k relative to \mathbb{Z}_+ is

$$\lim_{m \rightarrow \infty} \prod_{p \leq m} D_k^{(p)}(m) = \prod_p \left(1 - \frac{1}{p}\right)^{-k} . \quad (\text{A.6})$$

Meanwhile, to determine the contribution to dr from the ‘smoothed’ $\Lambda_{(k)}(n)$, follow the standard argument. There are $p - \nu_p(\mathcal{H}_k)$ residue classes mod p that n can occupy

that guarantee $n_{(k)}^k$ is not divisible by p . So the probability that $\Lambda_{(k)}(n)$ does not vanish given that n is a prime power is

$$\prod_{p < n_{(k)}^k} \left(1 - \frac{\nu_p(\mathcal{H}_k)}{p}\right). \quad (\text{A.7})$$

It follows that the integrator dr associated with the probability measure on \mathbf{r}_k is

$$dr = C_{(k)}(\mathbf{x}) dt := \prod_{p < x_{(k)}^k} \left(1 - \frac{\nu_p(\mathcal{H}_k)}{p}\right) \prod_p \left(1 - \frac{1}{p}\right)^{-k} dt \quad (\text{A.8})$$

where dt is integrator associated with the normalized Haar measure on \mathbb{R}_+ . Finally, define $C_{(k)}$ to be the asymptote

$$C_{(k)} := \lim_{x \rightarrow \infty} C_{(k)}(\mathbf{x}) = \prod_p \left(1 - \frac{\nu_p(\mathcal{H}_k)}{p}\right) \left(1 - \frac{1}{p}\right)^{-k}. \quad (\text{A.9})$$

References

- [1] G.H Hardy and J.E. Littlewood, Some problems of ‘Partitio numerorum’; III: On the expression of a number as a sum of primes, *Acta Math.*, **44**(1), (1923), 1–70.
- [2] B. Green and T. Tao, The primes contain arbitrarily long arithmetic progressions, *Ann. Math.*, **167**, (2008), 481–547
- [3] D.A. Goldston, J. Pintz, and C.Y. Yildirim, Primes in tuples I, *Ann. Math.*, **170**(2), (2009), 819–862.
- [4] J. Korevaar, Prime pairs and the zeta function, *Jour. Approx. Theor.*, **158**, (2009), 69–96.
- [5] Yitang Zhang, Bounded gaps between primes, *Ann. Math.*, **179**(3), (2014), 1121–1174.
- [6] J. LaChapelle, Exact Summatory Functions for Prime k -tuples, arXiv:math.NT/1406.5533 (2014).
- [7] J. LaChapelle, Evidence of a Gamma Distribution for Prime Numbers, arXiv:math.NT/1307.1717 (2013).
- [8] László Tóth, The Probability That k Positive Integers Are Pairwise Relatively Prime, *Fibonacci Quarterly* **40**(1), (2002), 13–18.